Landscape Ecol (2012) 27:253–266
DOI 10.1007/s10980-011-9701-4

RESEARCH ARTICLE

The influence of landscape characteristics and home-range size on the quantification of landscape-genetics relationships

Tabitha A. Graves · Tzeidle N. Wasserman · Milton Cezar Ribeiro · Erin L. Landguth · Stephen F. Spear · Niko Balkenhol · Colleen B. Higgins · Marie-Josée Fortin · Samuel A. Cushman · Lisette P. Waits

Received: 3 June 2011 / Accepted: 14 December 2011 / Published online: 3 January 2012

Abstract A common approach used to estimate landscape resistance involves comparing correlations of ecological and genetic distances calculated among individuals of a species. However, the location of sampled individuals may contain some degree of spatial uncertainty due to the natural variation of animals moving through their home range or measurement error in plant or animal locations. In this study, we evaluate the ways that spatial uncertainty, landscape characteristics, and genetic stochasticity interact to influence the strength and variability of conclusions about landscape-genetics relationships. We used a neutral landscape model to generate 45 landscapes composed of habitat and non-habitat, varying in percent habitat, aggregation, and structural connectivity (patch cohesion). We created true and alternate locations for 500 individuals, calculated ecological distances (least-cost paths), and simulated genetic distances among individuals. We compared correlations between ecological distances for true and alternate locations. We then simulated

Electronic supplementary material The online version of this article (doi:10.1007/s10980-011-9701-4) contains supplementary material, which is available to authorized users.

T. A. Graves (✉) · T. N. Wasserman
School of Forestry, Northern Arizona University, P.O. Box 15108, Flagstaff, AZ 86011, USA
e-mail: tabgra@yahoo.com

M. C. Ribeiro
Departamento de Ecologia, Universidade Estadual Paulista (UNESP), Av. 24A, 1515 Bela Vista, Rio Claro, SP 13506-900, Brazil

M. C. Ribeiro · C. B. Higgins
Department of Integrative Life Sciences, Virginia Commonwealth University, 1000 W. Cary Street, Richmond, VA 23284, USA

E. L. Landguth
Division of Biological Sciences, University of Montana, 32 Campus Drive, Missoula, MT 59812, USA

S. F. Spear
Orianne Society, 579 Highway 441 South, Clayton, GA 30525, USA

S. F. Spear · L. P. Waits
Department of Fish and Wildlife Resources, University of Idaho, P.O. Box 441136, Moscow, ID 83844, USA

N. Balkenhol
Department of Forest Zoology and Forest Conservation, University of Goettingen, Buesgenweg 3, 37077 Goettingen, Germany

M.-J. Fortin
Department of Ecology and Evolutionary Biology, University of Toronto, 25 Harbord St., Toronto, ON M5S 3G5, Canada

S. A. Cushman
Forest Service Research Station, US Forest Service, Flagstaff, AZ 86011, USA

genotypes at 15 neutral loci and investigated whether the same influences could be detected in simple Mantel tests and while controlling for the effects of isolation-by-distance using the partial Mantel test. Spatial uncertainty interacted with the percentage of habitat in the landscape, but led to only small reductions in correlations. Furthermore, the strongest correlations occurred with low percent habitat, high aggregation, and low to intermediate levels of cohesion. Overall genetic stochasticity was relatively low and was influenced by landscape characteristics.

Keywords Least cost · Habitat resistance · Fragmentation · Genetic structure · Sampling error · Aggregation · Cohesiveness · Connectivity · Gene flow · Isolation-by-resistance

Introduction

The degree of landscape resistance influences the movement of individual organisms and the genetic exchange among spatial localities (Murphy et al. 2008; Spear et al. 2010). Analyzing the genetic patterns that result from this exchange allows researchers to identify landscape features that enhance or diminish species movement, which in turn affects gene flow (Manel et al. 2003; Holderegger and Wagner 2008; Storfer et al. 2010). Landscape genetic studies have demonstrated how topographic and environmental gradients, habitat fragmentation, and barriers such as roads, streams and mountains create genetic structure within and among populations (e.g., Jacquemyn et al. 2004; Sork and Smouse 2004; McRae et al. 2005; Cushman et al. 2006; Ficetola et al. 2007; Dixo et al. 2009; Landguth et al. 2010; Murphy et al. 2010). Results from such studies are increasingly used to inform management and conservation applications, specifically with respect to corridor design (Epps et al. 2007; Braunisch et al. 2010; Segelbacher et al. 2010).

Among the most popular approaches for landscape genetic analyses is the statistical comparison of genetic and ecological distances (Storfer et al. 2010). In this approach, genetic distances are calculated among sampling units at the individual or population level using a diversity of metrics (Storfer et al. 2007, 2010). Ecological distance can be estimated by a number of approaches: Euclidean (straight-line), least-cost, least-corridor, or circuit-theory resistance distances (McRae 2006; Pinto and Keitt 2009; Rayfield et al. 2010). By regressing or correlating the genetic and ecological distances statistically, researchers can infer the effect of landscape-level features on genetic structure (Cushman et al. 2006; Balkenhol et al. 2009).

When individuals rather than populations are the unit of analysis, researchers typically use one location to represent each individual (Cushman et al. 2006; Schwartz et al. 2009; Braunisch et al. 2010; Shirk et al. 2010; Wasserman et al. 2010; Short Bull et al. 2011). Such a location may be the only sample for an individual, the center of multiple sample locations of an individual, or a location randomly chosen from multiple sample locations. Using just one geographic location to represent individuals in space may accurately represent the location of plants or sessile animals. Yet, when dealing with mobile animals, the use of a single location may introduce spatial uncertainty into landscape genetic analyses. The sampling location of highly vagile species may be far from central mating or reproduction areas, and therefore may not reflect environmental factors impacting gene flow. The influence of this potential spatial uncertainty on conclusions for landscape-genetics relationships likely depends on home-range size and specificity of breeding habitats. Many large mammals have extensive home ranges and pass through multiple habitat types, and thus may be sampled outside areas related to successful reproduction or survival (Lovari et al. 2008; Richard et al. 2008; Krofel et al. 2010). Although this is an obvious concern for most animal researchers, this problem cannot be easily addressed in an empirical study. Since researchers rarely know how well a specific sampling location represents the home range of a mobile animal, this potential issue is largely ignored in current studies of landscape genetics.

In addition to home-range size, the influence of spatial uncertainty may depend on landscape characteristics. Cushman et al. (in press) demonstrated that habitat area and fragmentation determine the significance of correlations between genetic and landscape distances. Landscapes with a high proportion of suitable habitat and low fragmentation are less likely to have detectable effects on genetic structures. Given this relationship, assessment of the influence of spatial uncertainty requires evaluation of the relative influence of landscape characteristics versus spatial uncertainty on landscape-genetics relationships.

We further hypothesized that increasing spatial uncertainty would increase the variability of

correlations in landscape-genetics relationships. Specifically, changes in individual locations could lead to changes in inter-individual ecological distances, which could change the strength of the correlation with genetic distances, potentially leading to different conclusions. The variation in locations also introduces variability into correlations between genetic and ecological distances.

Yet another influence on the strength and variability of correlations depends on genetic stochasticity (the random process of allele inheritance across generations). This genetic stochasticity adds variation to genetic distances between individuals, which could increase or decrease correlations. Finally, landscape characteristics could also influence the amount of genetic stochasticity. For instance, if the distance between habitat patches is near the limit of a species' mating or dispersal distance, gene flow may or may not occur, thus increasing the variation in genetic distances among individuals. Although real populations will only have one realization of this stochastic genetic process (and thus the variation cannot be measured empirically), by simulating multiple independent genetic datasets for each landscape scenario we can quantify the influence of all of these effects (spatial uncertainty, landscape characteristics, and genetic stochasticity) on the strength and variability of landscape-genetic conclusions.

In this study, we simulated landscape genetic datasets to provide a first assessment of whether variation in sampling location within an individual's home-range influences results in landscape genetic studies. We simulated individuals that move and reproduce across 45 landscapes with varying composition and configuration of habitat. We then altered the spatial location of individuals within five home ranges of increasing size. We evaluated the concordance of correlations obtained with the 'true' locations of individuals versus those obtained from five 'alternate' sampling points within the individual's home-range. We investigated how spatial uncertainty in individual locations affect (1) ecological distances between individuals and (2) correlations between genetic and ecological distances. In addition, we assessed the relative contributions of spatial uncertainty, landscape characteristics, and genetic stochasticity to the strength and variability of correlations used for landscape genetic inferences. We discuss the relevance of these findings for landscape genetic studies and suggest future avenues for research.

Methods

Simulating landscapes

We used a factorial study design to explore the way that spatial uncertainty, percent habitat, and fragmentation affect the strength of landscape-genetics relationships. For this, we used the neutral landscape model QRULE (Gardner 1999) to simulate binary landscape maps (habitat vs. non-habitat) with 512×512 pixels. QRULE controls fragmentation through the H parameter, which affects the aggregation of pixels into homogeneous patches. Higher values of H lead to higher levels of aggregation. We created landscapes that varied systematically across three levels of percent habitat (15, 35, and 55%) and aggregation ($H = 0.3$, 0.6, and 0.90; Fig. 1). We produced 5 replicates of this 9 combination factorial to assess variation and improve statistical power, yielding 45 total landscapes.

Simulating individuals

In each of the 45 landscapes, we randomly placed 500 individuals in habitat pixels to represent the true location of an individual (true locations; Fig. 1). To examine the effect of spatial uncertainty, we then placed five buffers equal to 10, 15, 20, 25, and 30% of the diameter of the landscape (equivalent to radii of 768, 1152, 1536, 1920, and 2304 m) around each true individual. Buffers represent different home range sizes which in essence correspond to different amounts of spatial uncertainty in sample locations. We created alternate locations for each individual by randomly placing a (x, y) location within each buffer (alternate locations; Fig. 2). Thus, for each of our 45 landscape maps, we had 500 true locations and 2,500 (500×5) alternate locations. Our buffers do not represent home ranges for any specific species, but instead (1) spatial uncertainty for any species with home ranges comprising 10–30% of the study area or (2) spatial error equal to 10–30% of the study area.

Calculation of pairwise distances between individuals

We calculated two-dimensional Euclidean (straight-line) distances between each pair of true individual locations for each landscape. Next we assigned a

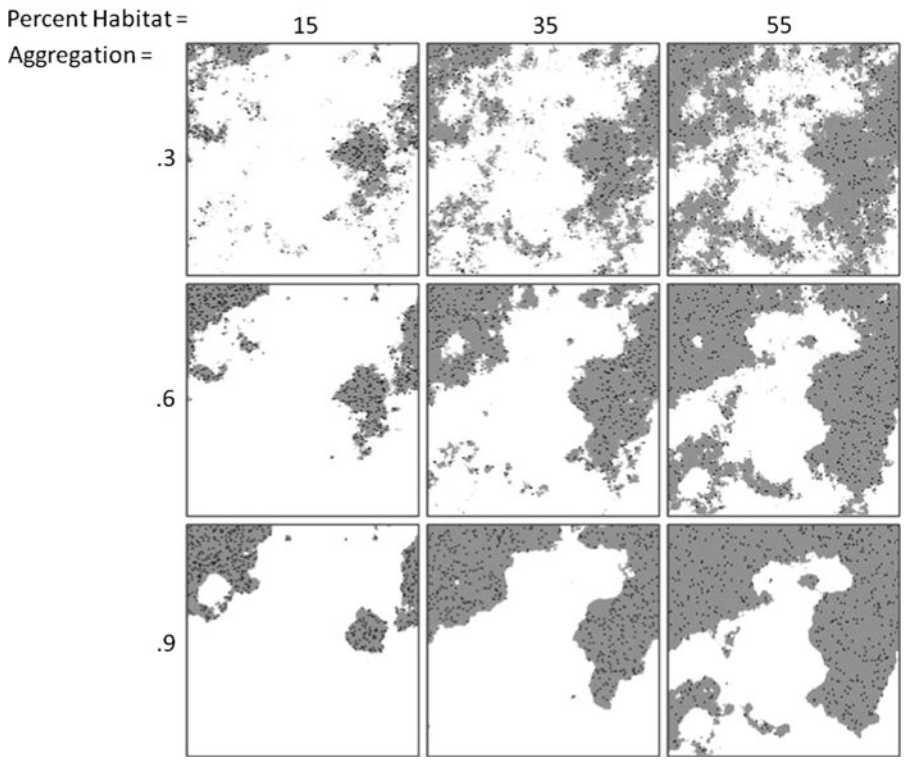

Fig. 1 A single replicate (9 landscapes) of our factorial design with true locations (*black dots*) in habitat (*gray areas*). We had 5 replicates for 45 total landscapes

resistance value of 20 to non-habitat landscape cells, compared to a resistance value of 1 for habitat cells, and calculated pair-wise ecological landscape distances between all true individuals and alternate individuals based on the accumulated least-cost path, using the COSTDISTANCE function in ArcGIS (ESRI 2008).

Simulating landscape-genetic relationships

We used CDPOP version 0.88 (Landguth and Cushman 2010) to simulate the processes of mating and dispersal as functions of the spatial patterns of habitat and non-habitat in these 45 simulated landscapes with true locations and pairwise landscape distances among true locations. CDPOP is an individual-based, spatially explicit, landscape genetics program that simulates birth, death, mating, and dispersal of individuals in complex landscapes as probabilistic functions of movement among the individuals (pairwise distances). We simulated gene flow among these true locations for 300 non-overlapping generations, long enough for genetic structure to stabilize (i.e., reach a spatial genetic equilibrium; Landguth et al. 2010). We generated genetic data for 15 loci, with 15 alleles that were initially randomly assigned among individuals (i.e., maximum allelic diversity). We used an inverse-square mating and dispersal probability function, with maximum dispersal distance of 50,000 m in uniformly suitable habitat. Reproduction was sexual, females could only mate a single time (without replacement) and males could mate with multiple females (with replacement). The number of offspring was based on a Poisson probability with a mean of 5, which ensured no immigration. For each of the 45 true landscape maps, we ran 10 Monte Carlo replicates to assess genetic stochastic variability. CDPOP produced matrices of pairwise genetic distances between all 500 simulated individuals based on proportion of shared alleles (Bowcock et al. 1994) at the end of our simulation.

Selection of landscape fragmentation metrics

In addition to percent habitat and aggregation, we wanted to examine landscape metrics that measured fragmentation and thus could potentially better predict the strength of landscape-genetic correlations. We selected landscape metrics based on past work, which assessed the strength and functional shape of the

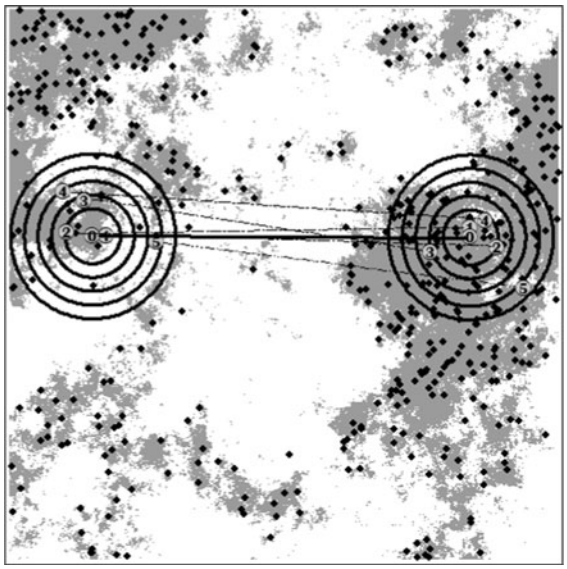

Fig. 2 Example of alternate point selection. *Black dots* show true locations. *Circles with numbers* show 2 true locations (0) and the alternate locations (1–5) chosen within buffers ranging from 10 to 35% of the diameter of the landscape. *Solid line* is Euclidean distance between true locations. *Dashed lines* show Euclidean distances between alternate locations. Least-cost paths between these pairs were used in analyses. Alternate locations could be anywhere within the outer buffer. The choice of alternate locations could lead to shorter (e.g. 3) or longer distances (e.g. 2), between a pair of individuals

relationship between a large number of landscape metrics and the H level in QRULE landscapes (Neel et al. 2004). Schumaker (1996) demonstrated that patch cohesion (hereafter, cohesion), which incorporates patch area and perimeter, very successfully predicted the ability of simulated owls to disperse in a landscape. Cushman et al. (in press) further found that cohesion strongly predicted gene flow. Their models also included a metric, clumpy (McGarigal et al. 2002), designed to approximate the aggregation created from the H level in QRULE. We therefore chose clumpy and cohesion as potential covariates. We calculated all metrics with FRAGSTATS (McGarigal et al. 2002).

Data analysis

To evaluate the influence of spatial uncertainty, landscape characteristics, and genetic stochasticity on correlations obtained in landscape-genetics studies, we carried out analyses for three different dependent variables. We used Mantel (Mantel 1967) and partial Mantel correlations (Smouse et al. 1986) as our dependent variables because this is the most common approach to assess landscape-genetics relationships. We then constructed different models that explain these three different dependent variables as a function of potential covariates (see below).

We calculated the Mantel's *r* correlation between ecological distances obtained from the 'true' individual locations, and the ecological distances obtained from the five alternate locations for each individual (True ~ Alternate). This correlation measures how closely ecological distances obtained from randomized locations match the 'true' ecological distances used to simulate the genetic patterns, so effects are not influenced by genetic stochasticity.

Then, we calculated Mantel's *r* between genetic distances and ecological distances obtained from alternate locations (Genetic ~ Ecological). This correlation quantifies the strength of landscape-genetics relationships resulting from our simulations.

Finally, we calculated partial Mantel's *r* between genetic distances and ecological distances, after accounting for the effects of Euclidean distances (Genetic ~ Ecological | Euclidean). This correlation measures the strength of the relationship between ecological and genetic distance after removing the influence of Euclidean distance on genetic distance. Genetic ~ Ecological and Genetic ~ Ecological | Euclidean are the primary approaches used to identify effects of landscape on genetic structure. In particular, recent simulation work demonstrated that partial Mantel tests have high power to correctly identify the drivers of genetic differentiation while rejecting incorrect alternative hypotheses in individual-based landscape genetics (Cushman and Landguth 2010).

Identifying covariate influence through generalized mixed models

To compare the relative influence of spatial uncertainty, landscape characteristics, and genetic stochasticity on the three Mantel correlation values described above, we modelled each correlation as a function of covariates. To do so, we used generalized linear mixed models (GLMM) and followed the procedure of Zuur et al. (2009) to identify the best model for each dependent variable. Each analysis incorporated three components: (1) random effects, (2) variance covariates, and (3) fixed effects.

The random effects component allowed us to (1) partition the variance resulting from genetic stochasticity versus landscape and individual configuration and (2) account for nuisance effects of using the same parameters to create multiple simulations. We had 10 genetic simulations for each landscape and set of individuals. Using landscape as a random effect allowed us to estimate the variation in correlations within a landscape due to genetic stochasticity. We had 5 replicates of the 9-factor combination of habitat amount and aggregation. Within a replicate, landscapes had similar configurations (Fig. 1), which we needed to account for to avoid pseudo-replication. To do so, we tested whether replicate was an important random effect. In addition to accounting for group correlations and allowing us to partition variance, random effects appropriately represent the landscapes and genetics we simulated, as a sample from a nearly infinite total we could have sampled.

The variance covariate component removed the assumption of homogeneity of variance and allowed us to model the influence of spatial uncertainty and landscape characteristics on the variability of our correlations. The fixed effects component models the strength of the influence of our spatial uncertainty and landscape characteristics on our dependent variable (correlations), in the same fashion as a typical linear regression.

Before we began modeling, we tested for linear correlations among explanatory variables, and eliminated highly correlated variables ($r > 0.9$) to avoid problems with collinearity. As the first step in finding the best model, we tested whether to include random effects for replicate or landscape in our model. We held the fixed effects component constant (most complex model) and compared models with (1) no random effects, (2) replicate-only random effects, (3) landscape-only random effects, and (4) landscape-nested-within-replicate random effects. We kept the random effect component from the model with the lowest akaike information criteria (AIC; Akaike 1973). AIC has two components: a term that assesses lack of fit and a penalty term for each additional parameter. The lowest score thus represents the most parsimonious model for the data, and minimizes both bias from underfitting and variance from overfitting (Burnham and Anderson 2002; Anderson et al. 2000).

Next, for the variance component, we modeled residual variance as a function of covariates. Holding the other two components constant (best random effects component and the most complex model for fixed effects), we found the best variance covariate component in two steps. First, we included one covariate in the variance component of the model and examined whether a multiplicative, exponential, power, or constant plus power relationship best described the form of variance with that covariate (Zuur et al. 2009). Then we compared all combinations (no variables to all variables) of the best model for each variance covariate. We kept the variance covariates from the model with the lowest AIC.

Finally, with the best random effect and variance covariate components included, we used backwards stepwise regression to identify the best fixed effect component of our model. Our most complex model included a quadratic term for cohesion, a two-way interaction term for spatial uncertainty $\times$ percent habitat, and a three-way interaction term for percent habitat $\times$ aggregation $\times$ cohesion, plus all main effect and two-way interactions within the three-way interaction (Supplement I). We examined standardized residuals to assess whether we adequately modeled heterogeneity of variance and histograms to assess normality of residuals.

We also calculated the proportion of significant partial Mantel's r for correlations of Genetic $\sim$ Ecological | Euclidean distances across levels of spatial uncertainty and percent habitat. All analyses were conducted in R (R Development Core Team 2010), using the packages ecodist for Mantel tests (Goslee and Urban 2007) and nlme for GLMMs (Pinheiro et al. 2011).

Results

Preliminary data exploration

Linear correlations among potential covariates ranged from $r = 0$ to $r = |0.56|$, except for the correlation between aggregation and clumpy ($r = 0.94$). Since clumpy is a landscape metric designed to closely approximate theoretical aggregation values, and aggregation was specifically set to simulate landscapes, we retained aggregation and removed clumpy from further analysis.

Model 1: True $\sim$ Alternate ecological distances

Mantel's r correlations between the true and alternate location ecological distances were all very high, with a

Table 1 Model coefficients from the best models explaining Mantel's r or partial Mantel's r. We had 225 Mantel's r correlations (45 landscapes × 5 alternate locations) for the model of True ~ Alternate ecological distances. After excluding 3 landscapes because the simulations did not meet objectives, we had 2,520 correlations (42 landscapes × 6 locations × 10 genetic simulations) for the models of Genetic ~ Ecological and Genetic ~ Ecological | Euclidean distances

Fixed Effects	True ~ Alternate Estimate (SE)	Genetic ~ Ecological Estimate (SE)	Genetic ~ Ecological | Euclidean Estimate (SE)
Spatial uncertainty	0.088 (0.021)	−0.052 (0.021)	0.062 (0.034)
Percent habitat	−0.065 (0.026)	−0.017 (0.002)	−0.009 (0.002)
Spatial uncertainty × Percent habitat	−0.013 (0.001)	−0.002 (0.001)	−0.003 (0.001)
Aggregation	−0.040 (0.012)	0.005 (0.001)	0.485 (0.227)
Cohesion	−0.006 (0.004)	21.2305 (9.891)	0.027 (0.103)
Cohesion^2		−0.108 (0.050)	
Aggregation × Cohesion	0.004 (0.001)		−0.005 (0.002)
Percent habitat × Aggregation	0.003 (0.001)		
Percent habitat × Cohesion	0.001 (<0.001)		
Percent habitat × Aggregation × Cohesion	−0.00003 (<0.001)		
Initial percent variance[a]			
Replicate	8.2		16.2
Landscape	91.8	92.9	73.9
Genetic		7.1	9.9
Final percent variance[b]			
Replicate	22	–	19.9
Landscape	78	100	80.1
Genetic		0	0

[a] Variances represent 'raw' variances, before variance covariates explaining the residual variation among genetic simulations are added to the model

[b] Variances represent final variances, after variance covariates explaining the residual variation among genetic simulations are added to the model

median of $r = 0.97$. Correlations always decreased with increasing spatial uncertainty, as expected. Spatial uncertainty, percent habitat, aggregation, and cohesion all significantly contributed to changes in the correlation between true and alternate location ecological distances (Table 1). Percent habitat and spatial uncertainty interacted. At low percent habitat (15%), predicted correlations changed only by 0.04, but the effect of spatial uncertainty increased at higher percent habitat (55%) to 0.125 (Fig. 3a). Highest correlations occurred with low percent habitat, high aggregation, and intermediate cohesion, which all interacted with each other. Breaking down the three-way interaction, as percent habitat increases, the change in correlation across cohesion becomes greater with increasing aggregation (Fig. 4a–c). At low habitat, cohesion has very little effect on Mantel's r across all levels of aggregation, but as habitat increases, cohesion and aggregation interact more, and the effect size of cohesion and aggregation increases as well. At maximum, effect size (change in predicted correlations) of these three variables was 0.129 (Fig. 4c). Thus, changes in spatial uncertainty or landscape characteristics could lead to similar decreases in correlations among true and alternate ecological distances.

The best model to explain the correlation between ecological distances based on true and alternate spatial locations included replicate random effects (Table 1), indicating that correlations between ecological distances based on true and alternate spatial locations for landscapes within a replicate were correlated with each other. The best variance structure included spatial uncertainty, percent habitat, and aggregation (Fig. 5). The variance of Mantel's r increased with power functions for spatial uncertainty and percent habitat, and exponentially for aggregation. Variance of correlations increased the most with percent habitat

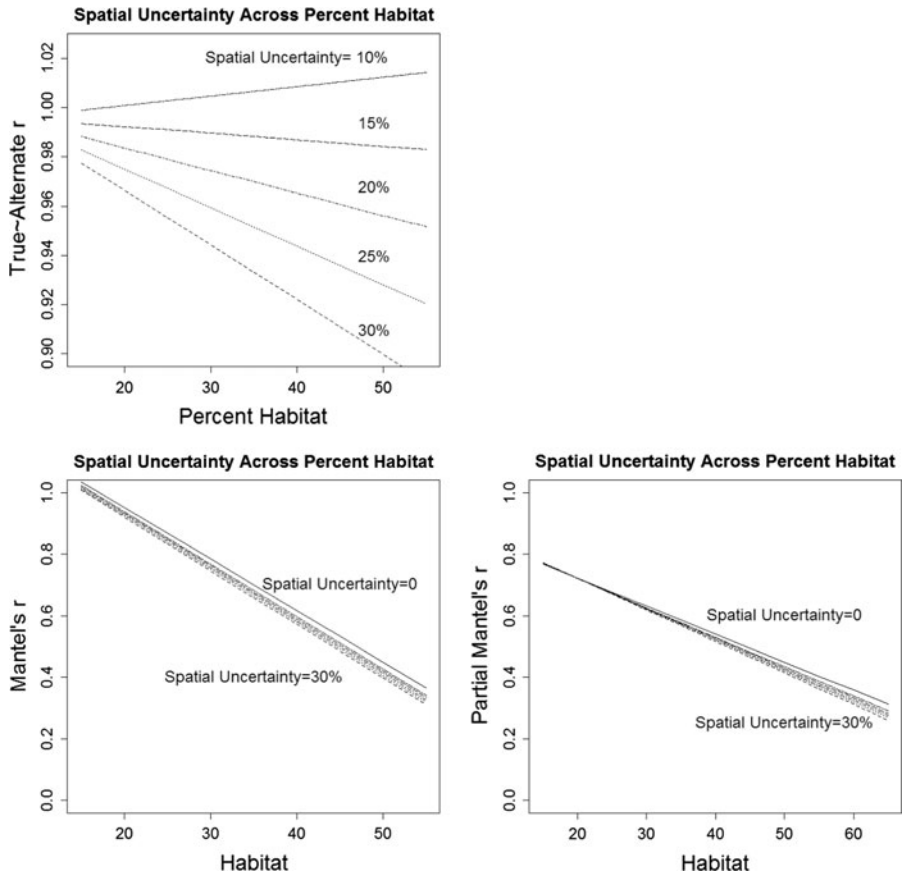

Fig. 3 Predicted influence of spatial uncertainty across percent habitat for **a** True ~ Alternate landscape distances, **b** Genetic ~ Landscape distances, and **c** Genetic ~ Landscape | Euclidean distances

(weights ranged from 0.01 to 0.17), but only slightly with aggregation (weights ranged from 0.0004 to 0.00007) and spatial uncertainty (weights ranged from 1.9×10^{-9} to 1.9×10^{-7}).

Model 2: Genetic ~ Ecological distance

Spatial uncertainty, percent habitat, aggregation, and cohesion all significantly influenced changes in the correlation between genetic and ecological distances (Table 1). Spatial uncertainty and percent habitat both have negative relationships with correlations between genetic and landscape distance. Spatial uncertainty and percent habitat interact, but the influence of percent habitat is much larger than the influence of spatial uncertainty (Fig. 3b). Correlations decrease more quickly at higher levels of percent habitat and higher spatial uncertainty. Mantel's r increases slightly with aggregation of habitat (Fig. 4d). Cohesion has a quadratic relationship with correlations between genetic and landscape distance (Fig. 4e). Correlations of genetic and ecological distances increase slightly up to cohesion ~98 and then sharply decrease. The highest Mantel's r again occurred with low percent habitat, high aggregation, and intermediate cohesion.

The model with landscape-only random effects was best supported (Table 1). Genetic stochasticity accounted for only 7.1% of the variation remaining after inclusion of spatial uncertainty and landscape characteristics as fixed effects. Furthermore, after modeling the genetic stochasticity as a function of variance covariates, less than 1% of the variation was due to genetic stochasticity alone (Table 1). The best variance covariate model included an interaction between percent habitat and aggregation. Genetic stochasticity increased across aggregation, slightly at 15% habitat and strongly at 55% habitat (Fig. 5b, c).

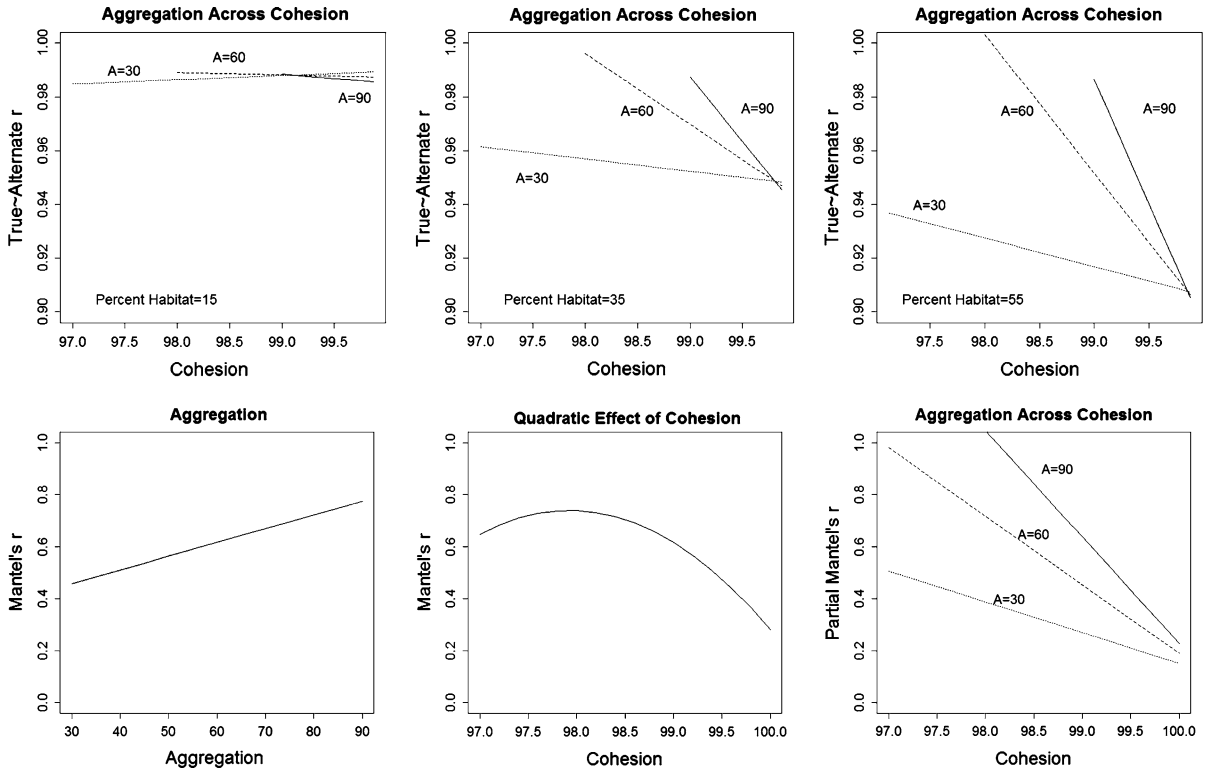

Fig. 4 Predicted influence of aggregation and cohesion. **a**–**c** Aggregation across cohesion for True ~ Alternate landscape distances (by habitat), **d** Linear influence of aggregation for Genetic ~ Ecological distances, **e** Quadratic influence of cohesion for Genetic ~ Ecological distances, and **f** Aggregation across cohesion for Genetic ~ Ecological | Euclidean distances

The largest variance weights are 0.0043. Standardized residuals from the final model are homogeneous and normal.

Model 3: Genetic ~ Ecological | Euclidean distances

As with the other response variables, spatial uncertainty and our three landscape characteristic covariates all influenced the strength of correlations between genetic and ecological distances, even with the effect of Euclidean distance partialled out. Percent habitat and spatial uncertainty interact, as they did for our other response variables. Partial Mantel's *r* correlations always decrease with increases in percent habitat but decrease more quickly with higher spatial uncertainty (Fig. 3c). Overall the influence of spatial uncertainty on strength of genetic structure is very small with maximum differences in partial Mantel's *r* of less than 0.04. Percent habitat has a relatively large effect size, with predicted partial Mantel's *r* ranging from 0.40 with 55% habitat to 0.77 with 15% habitat (Fig. 3c). Aggregation and cohesion also interact (Fig. 4f). Mantel's *r* increases with aggregation and decreases with cohesion, but decreases more quickly at higher aggregation. At aggregation of 60, partial Mantel's *r* ranged from 0.95 at our minimum (97.13) to 0.22 at our maximum (99.87) level of cohesion, indicating that cohesion can have a strong influence on the strength of genetic inferences. Aggregation has a slightly lower effect size, even with cohesion of 98.2 (highest effect size of aggregation). Partial Mantel's *r* ranged from 0.36 at our minimum (30) to 0.97 at our maximum (90) aggregation. Highest partial Mantel's *r* occurred with low percent habitat, high aggregation, and low cohesion, as they did for our other response variables. Standardized residuals from the final model are homogeneous and normal.

Inclusion of both replicate and landscape as random effects was supported. Prior to the addition of variance covariates and given the inclusion of fixed effects

covariates, genetic stochasticity accounted for 9.9% of the variation (Table 1). After adding variance covariates, variation among genetic simulations accounted for <1% of the variation, and most (>80%) of the remaining unexplained variation was among landscapes.

Variance decreased slightly with spatial uncertainty and percent habitat and increased slightly with aggregation and cohesion (Fig. 5). All variance weights were ≤0.000008, a very low effect size.

Significance tests for partial Mantel's r:
Genetic ~ Ecological | Euclidean distances

As with the strength of Mantel's *r* correlations, the proportion of significant correlations had an inverse relationship with percent habitat. Across spatial uncertainty, the proportion of significant partial Mantel's *r* correlations varied little (Table 2).

Discussion

Although many studies now incorporate landscape genetics into conservation planning, researchers have simply used the available samples without knowing the degree to which spatial uncertainty, landscape configuration, and genetic stochasticity affect the strength of genetic inferences. Because simulations provide multiple realizations of the stochastic genetic process, simulation modeling provides a platform to observe a greater variety of situations than is possible in empirical studies and allows explicit control over pattern-process relationships (Epperson et al. 2010). This allows researchers to investigate hypotheses about varying organism characteristics, population size, dispersal ability, the influence of spatial factors, and their interactions that are difficult to investigate directly in the field (Wasserman et al. 2011). The use of simulations allowed us to explore the effects of

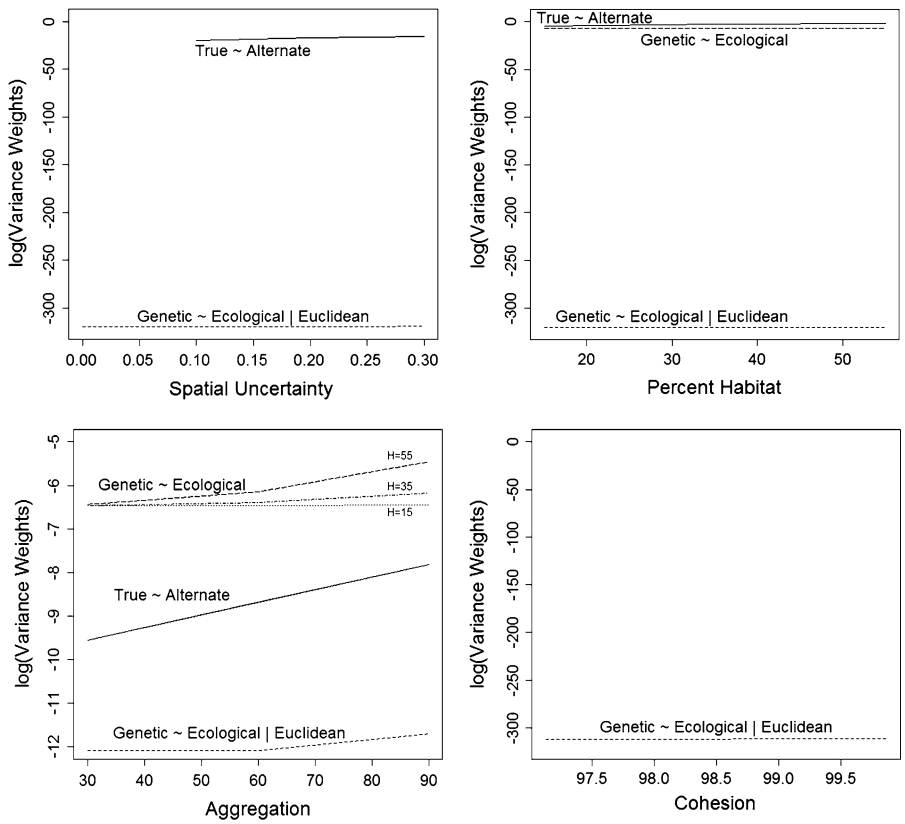

Fig. 5 Log of variance weights for Spatial uncertainty, Percent habitat, Aggregation, and Cohesion covariates. *Solid lines* represent variance covariate weights for the True ~ Alternate response variable. Intermediate *dashed lines* are Genetic ~ Landscape distances. *Short dashed lines* are Genetic ~ Landscape | Euclidean distances. Weights for True ~ Alternate are generally largest, because they represent variation at the landscape level, while weights for Genetic ~ Landscape|Distance represent only variation at the level of the genetic simulation

Table 2 Percentage of significant genetic relationships from partial Mantel's *r* for Genetic ~ Landscape | Euclidean distances

Percent habitat	Spatial uncertainty					
	0	0.1	0.15	0.2	0.25	0.3
15	0.94	0.95	0.94	0.95	0.93	0.94
35	0.99	0.99	0.99	0.99	0.99	0.97
55	0.71	0.73	0.73	0.71	0.71	0.70

landscape configuration and spatial uncertainty while holding other factors constant. We were also able to assess genetic stochasticity directly using simulated data. While our simulations represent rather simplified landscapes, did not evaluate demographic stochasticity, and included several assumptions about the dispersal and mating processes, our results provide some initial insights into the influence of sampling an individual at only a single location and the influence of landscape on landscape genetic inferences. Our results should also apply when obtaining samples of plants or sedentary animals that have location errors. Generally, our results provide reassurance to landscape genetics researchers with studies that have (1) only sampled an individual at one location within a home-range of unknown size and (2) the variable of interest can be simplified to habitat versus non-habitat.

Interplay of spatial uncertainty, landscape characteristics, and genetic stochasticity

As expected, increasing spatial uncertainty decreased Mantel's *r* correlations between true and alternate ecological distances. However, spatial uncertainty had only a very small effect on the strength of genetic structure (changes in $r < 0.05$) measured with Mantel's and partial Mantel's *r* for genetic and ecological distances, at least with the resistance value fixed at 20. Variance of Mantel's *r* correlations between true and alternate distances increased with spatial uncertainty, as expected, however, the variance of correlations between genetic and ecological distances was negligible, though significant, at this resistance value.

Percent habitat consistently had a relatively large effect size. Aggregation and cohesion also had substantial effect sizes, which interacted, making simple interpretations difficult. At high levels of cohesion, equivalent to a highly connected landscape, which only occurred under high aggregation, only large fragments of habitat existed, our populations approached panmixia, and little genetic structure existed. At low cohesion, equivalent to a less connected landscape, changes in aggregation (60) led to large changes in partial Mantel's *r*. At low levels of cohesion, increases in aggregation led to fewer 'stepping stones' for gene flow, and thus greater genetic structure. Because percent aggregation and cohesion interact, ranking effect size of these variables is only meaningful given a constant value of the other variable.

Caveats and further research

Our ability to detect interaction effects of landscape composition and configuration, as measured by percent habitat, aggregation, and cohesion decreased after the addition of genetic simulations (our Genetic ~ Ecological models), likely because landscape characteristics influenced not only ecological distances (Model 1), but also genetic distances (Models 2, 3). Greater genetic stochasticity existed in some landscapes than others, which we modeled using variance covariates (reducing residual variation in the genetic component to almost 0). Because our fixed effect covariates did not fully explain the variation from different landscape configurations, we believe that investigation of additional landscape covariates may provide additional insight. Our findings provide insight to conservation practitioners who are using estimates of resistances determined through causal modeling (Cushman et al. 2006; Shirk et al. 2010; Wasserman et al. 2010). Although we have not found a method to calculate uncertainty for resistances determined through causal modeling, knowing that uncertainty depends on the landscape characteristics can give insight into conservation planning efforts using resulting resistance estimates. For example, Short Bull et al. (2011) found that detectability of the influences of roads, elevation, and forest cover on genetic differentiation in American black bear was dependent on whether these features were structured in such a way as to limit movement and gene flow in the particular landscape.

Three situations exist in which spatial uncertainty could have a greater influence on genetic structure inferences than we found based on our simulations. First, higher landscape resistances could lead to greater variation in the strength of genetic structure. In a brief survey of published studies from natural

systems, optimized resistance values ranged from 1 to 10,000 (Epps et al. 2005; Cushman et al. 2006; Epps et al. 2007; Shirk et al. 2010; see Spear et al. 2010 for a list of these kinds of studies), considerably higher than the resistance (20) we tested. More work is necessary to identify whether genetic inferences are robust to random spatial uncertainty with larger resistance values or multiple habitat types, with varying or high resistance values (Rayfield et al. 2010). Second, if spatial errors are biased, such that a large number of spatial errors are inside a habitat with high resistance or across a barrier with very high resistance, the resistance of that habitat or barrier could be underestimated. This could be a problem when GPS or human error misidentifies locations in a biased way, such as might occur in heavily forested areas. A description of the type of habitat and direction relative to barriers could be used to check and reduce this type of error. Biased errors could also occur for plants if they are long-lived and the habitat becomes more resistant to gene flow around them. For example if a species that requires light to regenerate becomes overgrown by trees, this time lag could lead to an underestimate of the resistance of the forest. These effects should only change inferences when the same kind and direction of errors contaminate a high proportion of samples. Third, by using least-cost paths as our landscape distance metric, we assume that the individual knows the landscape and disperses ideally along a single best (least-cost) path. We do not know whether our results would hold with individuals that have less knowledge of the landscape and disperse via a random walk or partial random walk, which can be approximated with circuit theory (McRae 2006) or least-cost corridors (Pinto and Keitt 2009). Most studies in landscape genetics have relied on the use of Mantel and partial Mantel testing (Storfer et al. 2010). Future work should explore whether other distance-based methods, such as distance-based redundancy analysis (see Legendre and Fortin 2010), are more sensitive to the influence of spatial uncertainty. More work is needed to understand the influence of our assumptions (i.e., simplified landscapes, no demographic stochasticity, and small set of mating and dispersal parameters) on landscape genetic results.

Few researchers have explored how variation in the pattern of landscape mosaics affects genetic differentiation. Bruggeman et al. (2010) used simulation modeling to quantify the influence of patch size and patch isolation on abundance, effective population size, and Fst in red-cockaded woodpecker. Their work is one of the first to provide an explicit link between population genetic processes, habitat area, and critical thresholds of fragmentation affecting those processes. Their results suggest that population genetic structure is more strongly affected by habitat fragmentation than habitat patch size.

Cushman et al. (in press) quantified the relative importance of habitat area and configuration on genetic differentiation across broad gradients of the extent and fragmentation of habitat. They used spatially explicit, individual-based simulation modeling to quantify the effects of habitat area, fragmentation, and the contrast in resistance between habitat and non-habitat on the apparent strength and statistical detectability of landscape genetic relationships. They found that cohesion had the strongest ability to predict the magnitude of genetic differentiation, which is consistent with our findings that cohesion has a very large effect on genetic differentiation, although we found that the effect size is mediated by the level of aggregation. Cushman et al. (in press) also found that when habitat area is very high or habitat fragmentation is very low landscape structure does not limit gene flow, and landscape genetic effects are often not detectable. This is also consistent with our findings that lowest genetic differentiation occurred with high percent habitat, low aggregation, and high cohesion. Our work builds on Cushman et al. (in press) by examining the influence of spatial uncertainty, using a higher level of resistance (20 vs. 1–16), testing the theoretical value of aggregation (H) rather than aggregation index or clumpy, and our focus on the strength of genetic inferences rather than detection of landscape influences on genetic differentiation. Our results indicate that at least in landscapes with relatively low resistance and randomly distributed spatial error, the spatial uncertainty has little influence on inferences of genetic structure. Percent habitat, cohesion, and aggregation can all drive the strength of the correlations of genetic and landscape distances, although this relationship is relatively complex (Fortin et al. 2003; Neel et al. 2004; Bruggeman et al. 2010; Cushman et al. in press). The highest correlations consistently occurred with low percent habitat, high aggregation, and low-intermediate cohesion. This is consistent with expectations that genetic structure will be lower and thus more difficult to detect in landscapes

with relatively intact habitat. Most of the unexplained variation remained at the landscape level, so further investigation into other landscape metrics may prove fruitful.

Acknowledgments We thank Brent Burch, Kevin McGarigal, and John Citta for useful discussions on modeling random effects. This study resulted from a distributed graduate seminar (developing best practices for testing landscape effects on gene flow), conducted through the National Center for Ecological Analysis and Synthesis (NCEAS), a center funded by the National Science Foundation Grant #EF-0553768, the University of California, Santa Barbara, and the State of California. We thank Carisa Stansbury and Rodrigo Cisneros for assistance with preliminary simulations. This was a truly collaborative project. Contributions of each coauthor are listed in Supplement II. We appreciate our individual sources of support, including scholarship, fellowship and research assistantship providers. We also thank 2 anonymous reviewers.

References

Akaike H (1973) Information theory as an extension of the maximum likelihood principle. In: Petrov BN, Csaki F (eds) Second international symposium on information theory. Akademiai Kiado, Budapest, pp 267–281

Anderson DR, Burnham KP, Thompson WL (2000) Null hypothesis testing: problems, prevalence, and an alternative. J Wildl Manag 64(4):912–923

Balkenhol N, Gugerli F, Cushman SA, Waits LP, Coulon A, Arntzen JW, Holderegger R, Wagner H (2009) Identifying future research needs in landscape genetics: where to from here? Landscape Ecol 24:455–463

Bowcock AM, Ruiz-Linares A, Tomfohrde J, Minch E, Kidd JR, Cavalli-Sforza LL (1994) High resolution of human evolutionary trees with polymorphic micorsatellites. Nature 368:455–457

Braunisch V, Hirzel A, Segelbacher G (2010) Modelling functional landscape connectivity from genetic population structure: a new spatially explicit approach. Mol Ecol 19:3664–3678

Bruggeman DJ, Wiegand T, Fernandez N (2010) The relative effects of habitat loss and fragmentation on population genetic structure. Mol Ecol 19:3691–3697

Burnham KP, Anderson DR (2002) Model selection and multimodel inference: a practical information-theoretic approach. Springer, New York

Cushman SA, Landguth EL (2010) Spurious correlations and inference in landscape genetics. Mol Ecol 19:3592–3602

Cushman SA, McKelvey KS, Hayden J, Schwartz MK (2006) Gene flow in complex landscapes: testing multiple hypotheses with causal modeling. American Naturalist 168:486–499

Cushman SA, Shirk A, Landguth EL (2012) Separating the effects of habitat area, fragmentation and matrix resistance on genetic differentiation in complex landscapes. Landscape Ecol. doi:10.1007/s10980-011-9693-0

Dixo M, Metzger JP, Morgante JS, Zamudio KR (2009) Habitat fragmentation reduces genetic diversity and connectivity among toad populations in the Brazilian Atlantic Coastal Forest. Biol Conserv 142:1560–1569

Epperson BK, McRae B, Scribner K, Cushman SA, Rosenberg MS, Fortin M-J, James PMA, Murphy M, Manel S, Legendre P, Dale MRT (2010) Utility of computer simulations in landscape genetics. Mol Ecol 19:3540–3564

Epps CW, Palsboll PJ, Wehausen JD, Roderck GK, Ramey RR II, McCullough DR (2005) Highways block gene flow and cause a rapid decline in genetic diversity of desert bighorn sheep. Ecol Lett 8:1029–1038

Epps CW, Wehausen JD, Bleich VC, Torres SG, Brashares JS (2007) Optimizing dispersal and corridor models using landscape genetics. J Appl Ecol 44:714–724

ESRI (1999–2008) Environmental System Research Institute, Redlands

Ficetola GF, Garner TWJ, Biernard FD (2007) Genetic diversity, but not hatching success, is jointly affected by postglacial colonization and isolation in the threatened frog, Rana latastei. Mol Ecol 9:1787–1797

Fortin M-J, Boots B, Csillag F, Remmel TK (2003) On the role of spatial stochastic models in understanding landscape indices in ecology. Oikos 102:203–212

Gardner RH (1999) QRULE: a program for the generation of random maps and the analysis of spatial patterns. In: Klopatek JM, Gardner RH (eds) Landscape ecological analysis: issues and applications. Springer-Verlag, New York, pp 280–303

Goslee SC, Urban DL (2007) The ecodist package for dissimilarity-based analysis of ecological data. J Stat Softw 22:1–19

Holderegger R, Wagner HH (2008) Landscape genetics. Bioscience 3:199–207

Jacquemyn H, Honnay O, Galbusera P, Roldan-Ruiz I (2004) Genetic structure of the forest herb *Primula elatior* in a changing landscape. Mol Ecol 13:211–219

Krofel M, Filacorda A, Jerina K (2010) Mating-related movements of male brown bears on the periphery of an expanding population. Ursus 21:23–29

Landguth EL, Cushman SA (2010) CDPOP: an individual-based, cost-distance spatial population genetics model. Mol Ecol Resour 10:156–161

Landguth EL, Cushman SA, Schwartz MK, McKelvey KS, Murphy M, Luikarts G (2010) Quantifying the lag time to detect barriers in landscape genetics. Mol Ecol 19: 4179–4191

Legendre P, Fortin M-J (2010) Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. Mol Ecol Resour 10:831–844

Lovari S, Barolommei P, Meschi F, Pezzo F (2008) Going out to mate: excursion behaviour of female roe deer. Ethology 114:886–896

Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. Trends Ecol Evol 18:189–197

Mantel N (1967) The detection of disease clustering and a generalized regression approach. Cancer Res 27:209–220

McGarigal K, Cushman SA, Neel MC, Ene E (2002) FRAGSTATS: spatial pattern analysis program for categorical

maps. Computer software program produced by the authors at the University of Massachusetts, Amherst

McRae BH (2006) Isolation by resistance. Evolution 60: 1551–1561

McRae BH, Beier P, Dewald LE, Huynh LY, Keim P (2005) Habitat barriers limit gene flow and illuminate historical events in a wide-ranging carnivore, the American puma. Mol Ecol 14:1965–1977

Murphy MA, Evans JS, Cushman SA, Storfer A (2008) Representing genetic variation as continuous surfaces: an approach for identifying spatial dependency in landscape genetic studies. Ecography 31:685–697

Murphy MA, Evans JS, Storfer A (2010) Quantifying Bufo boreas connectivity in Yellowstone National Park with landscape genetics. Ecology 91:261–262

Neel MC, Cushman SA, McGarigal K (2004) Behavior and stability of landscape metrics across controlled gradients of landscape structure. Landscape Ecol 19:435–455

Pinheiro J, Bates D, DebRoy S, Sarkar D, The R Development Core Team (2011) Nlme: linear and nonlinear mixed effects models. R package version 3:1–100

Pinto N, Keitt TH (2009) Beyond the least-cost path: evaluating corridor redundancy with a graph theoretic approach. Landscape Ecol 24:253–266

R Development Core Team (2010). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org

Rayfield B, Fall A, Fortin M-J (2010) The sensitivity of least-cost habitat graphs to relative cost surface values. Landscape Ecol 25:519–532

Richard E, Morellet N, Cargnelutti B, Angibault JM, Vanpe C, Hewison AJM (2008) Ranging behaviour and excursions of female roe deer during the rut. Behav Process 79:28–35

Schumaker NH (1996) Using landscape indices to predict habitat connectivity. Ecology 77:1210–1225

Schwartz MK, Copeland JP, Anderson NJ, Squires JR, Inman RM, McKelvey KS, Pilgrim KL, Waits LP, Cushman SA (2009) Wolverine gene flow across a narrow climatic niche. Ecology 90:3222–3232

Segelbacher G, Cushman SA, Epperson BK, Fortin M-J, Francois O, Hardy DJ, Holderegger R, Taberlet P, Waits LP, Manel S (2010) Applications of landscape genetics in conservation biology: concepts and challenges. Conserv Genet 11:375–385

Shirk AJ, Wallin DO, Cushman SA, Rice CG, Wahrheit KA (2010) Inferring landscape effects on gene flow: a new model selection framework. Mol Ecol 19:3603–3619

Short Bull RA, Cushman SA, Mace R, Chilton T, Kendall KC, Landguth EL, Schwartz MK, McKelvey KS, Allendorf FW, Luikart G (2011) Why replication is important in landscape genetics: American black bear in the Rocky Mountains. Mol Ecol 20:1092–1107

Smouse PE, Long JC, Sokal RR (1986) Multiple regression and correlation extensions of the Mantel test of matrix correspondence. Syst Zool 35:627–632

Sork VL, Smouse PE (2004) Genetic analysis of landscape connectivity in tree populations. Landscape Ecol 21:821–836

Spear SF, Balkenhol N, Fortin M-J, Mcrae BH, Scribner K (2010) Use of resistance surfaces for landscape genetic studies: considerations for parameterization and analysis. Mol Ecol 19:3576–3591

Storfer A, Murphy M, Evans JS, Goldberg CS, Robinson S, Spear SF, Dezzani R, Delmelle E, Vierling L, Waits LP (2007) Putting the "landscape" in landscape genetics. Heredity 98:128–142

Storfer A, Murphy M, Spear S, Holderegger R, Waits L (2010) Landscape genetics: where are we now? Mol Ecol 19:3496–3514

Wasserman TN, Cushman SA, Schwartz MK, Wallin DO (2010) Spatial scaling and model inference in landscape genetics: *Martes Americana* in northern Idaho. Landscape Ecol 25:1601–1612

Wasserman TN, Cushman SA, Shirk AS, Landguth EL, Littell JS (2011) Simulating the effects of climate change on population connectivity of American marten (*Martes Americana*) in the northern rocky Mountains, USA. Landscape Ecol. doi:10.1007/s10980-011-9653-8)

Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM (2009) Mixed effects models and extensions in ecology with R. Springer, New York